Trend Analysis of Ferry Travel Patterns for Sustainable Transport Planning in Brisbane, Queensland, Australia

Roy Zhu¹, SangHyung Ahn¹, Michelle Robertson¹, Linda Lin¹, Summer Legambi² and Zachary Zweig²

¹School of Civil Engineering, The University of Queensland, Queensland 4072 Australia
²A. James Clark School of Engineering, The University of Maryland, Maryland 20742 USA

Email for correspondence: roy.zhu@uq.net.au

Abstract

Many cities across the world are facing challenges from unsustainable travel patterns including growing traffic congestion, overcrowding on public transport and aging transport infrastructure. The city of Brisbane, Australia continues to face extensive population and economic growth, which adds further pressure onto its existing transportation network and thus, the need to implement effective sustainable strategies becomes increasingly important. Ferry transit offers an opportunity for Brisbane, as it runs parallel with many of its major road corridors whilst connecting key areas of the city such as the central business districts, major tertiary education institutions and recreational areas. The system can also be used to bypass disruption on land based infrastructure such as road congestion. However, a better understanding of Brisbane’s ferry network is crucial to determine its potential for facilitating transit for regular commuters. Smart-card data is emerging as a source of insight into the performance of a city’s transport network. Past researchers have used Brisbane’s Go Card data to investigate broad, system-level trends in ferry usage. However, there is little existing literature which conducted microscopic analysis of ferry passenger travel patterns. This paper aims to adopt a more human-centric approach for smart card data analysis, with a focus on developing new visualisations to assist in understanding public transportation networks. Ultimately, three plots were proposed to assist in depicting terminal performance. The first was a box-and-whisker plot for visualising variations in demand throughout the day using a month of smart card data. The second plot was developed using a day of smart card data, and depicts the dwell time duration of each ferry due to individual passengers alighting and boarding. Lastly, a third plot was created to depict dwell time duration due to a specific quantity of alighting, boarding and total passengers for a particular ferry terminal. These plots begin to highlight inefficiencies prominent in the system, which could provide means for future research.

1. Introduction

In a predominantly urban world, the ability to reorient transportation infrastructure to allow people ease of access to move within cities to jobs, services and key facilities is a critical driver of sustainability. Brisbane is Australia’s third largest city and is forecasted to face significant population growth coupled with economic and employment growth within the next few years (Brisbane City Council, 2017). These trends emphasise the increasing demand and pressure on existing infrastructure and transport services in the region. Such transport challenges can no longer be resolved by building additional roads that funnel increasing traffic into an already congested urban core (Brisbane City Council, 2017). Therefore, it is crucial for Brisbane to commit to long-term planning for emerging issues as opposed to employing short-term solutions
One potential approach is the utilisation of non-road based transportation systems. In Brisbane, the CityCat and related cross-river ferries are an integral part of the transport network for residents and commuters (Moore, 2011) in addition to bus and rail. Ferry services generally hold a significant role in urban transportation systems across the world as they provide an alternative transit mode which eliminates the need for large investments into infrastructure that may not be economically or environmentally feasible. “Think globally, act locally” has become an increasingly new viewpoint in developing effective sustainable strategies (QCOSS, 2013). From a transport perspective, human-centred design approach is emphasised for solving problems relating to sustainability. Subsequently, it is critical to develop a clear understanding of commuter travel behaviour and habits.

The relationship between commuting patterns and land use has drawn increasing research interests to better understand and alleviate urban mobility related issues. Many of these commuting patterns are becoming accessible from the analysis of smart card data. Previous research undertaken by Soltani et al. in 2015 first proposed the use of smart card data for macroscopic analysis of travel behavior patterns within the ferry system (Soltani et al., 2015). The aim of the research underlying this paper was to build upon the work of Soltani et al. through analysing one slice of 6-month data, ranging from November 2015 to November 2016. However, this study adopts a more microscopic approach in attempt to extend knowledge on the information which can potentially be extracted from smart card data. Specifically, it seeks to introduce visualisation tools for modelling passenger behavior at ferry terminals.

The structure of this paper is as follows: First, we conduct literature review and start with providing background information for Brisbane’s ferry network. After discussing methodologies that used to analyse and visualise smart card data collected in April 2016. We then discuss about the results of the data analysis. We close the paper by briefly suggesting direction for future research and by providing a summarizing conclusion.

2. Review of Smart Card Data Analysis

Public transit agencies worldwide are increasingly adopting smart card fare collection systems. Whilst offering increased convenience and boarding efficiency, smart card usage can also provide vast amounts of data including route information, boarding and alighting information. Thus far, research outcomes from smart card data analysis have been classified as being strategic, tactical or operational level. Strategic-level studies focus on demand forecasting, long-term network planning and commuter behaviour analysis. Tactical-level studies relate to trip patterns and schedule adjustment. Operational-level studies investigate real-time performance, and smart card system operations. (Pelletier, Trepanier, & Morency, 2011)

The data available from smart card systems provides substantially more information than other methods of data collection, providing numerous opportunities for macroscopic levels of analysis (Agard, Morency, & Trepanier, 2006). For strategic-level and tactical-level studies, the timestamps and geospatial information often provided by smart card data allows researchers to analyse commuter travel patterns in a city, identifying desirelines and describing their variation with time. A study targeting both approaches was Soltani et al.’s study on Brisbane ferry usage. Their study allowed for a visualisation and discussion of (1) monthly variations (November 2012 to April 2013) and daily variations in CityCat usage (Monday to Sunday), (2) variation in CityCat transaction count throughout the day, (3) the demand of specific ferry terminals as origins and destinations, (4) the proportion of journeys which included intermodal transit (e.g. bus-to-ferry, ferry-to-train, etc. connections), (5) the proportion of ferry trips which were simply cross-river relative to longer, linear trips up- or downriver and (6) the frequency of ferry usage by passengers.

Unfortunately, whilst the scale and complexity of smart card data poses an opportunity for researchers, it also creates limitations and problems. Amongst issues with data validation, achieving precision when obtaining desired operational information (vehicle number, route
number, direction) remains a persistent challenge (Pelletier, Trepanier, & Morency, 2011). Although these issues are not significant at higher levels of analysis due to the quantity of data involved, the outcomes from macroscopic analysis are also limited in their applicability for decision-making for planners and operators at microscopic levels.

In response, Pelletier et al. proposes that new modelling and analytical methods are required, as classical models are unable to be applied to such detailed levels of resolution. Operational-level studies have begun this process, with some focusing on the development of performance indicators. These include investigating schedule adherence for services on each individual run, route or day (Trepanier, Morency, & Agard, 2009). Zhou et al. developed a model for calculating bus arrival times based on passenger swiping behaviours. This would allow for accurate timekeeping for passenger information systems, whilst provide a means of evaluating bus operation efficiency (Zhou et al., 2017). As such, while this paper will further the work conducted by Soltani et al. analysing Go Card Data, its primary purpose is to contribute towards operational-level studies by introducing visualisation tools which increase the amount of information and insights extractable from smart card data.

3. Research Context and Literature Review

3.1. Brisbane’s Ferry Network

The strategic intent of the Brisbane ferries was initially focused on providing a tourism opportunity. However they became a crucial part of the city’s transport network for commuters due to the opportunities provided by Brisbane River’s ability to allow transit to key facilities without the interference of road-based congestion (Sipe & Burke, 2011). Brisbane’s ferry transport network consists of three passenger only services; CityCat catamaran ferries, CityHopper inner city ferries and the Cross River Ferry. Their details are summarised in Table 1 below. These ferry services are integrated into Brisbane’s transport network with passengers being able to access park-and-ride facilities, pedestrian and cyclist facilities, busses and rail from ferry terminals.

<table>
<thead>
<tr>
<th>Vehicle</th>
<th>Number of vessels as of 2016</th>
<th>Passenger capacity as of 2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>CityCat Total: 30</td>
<td>First generation: 8</td>
<td>First generation: 149</td>
</tr>
<tr>
<td></td>
<td>Second generation: 6</td>
<td>Second generation: 162</td>
</tr>
<tr>
<td></td>
<td>Third generation: 7</td>
<td>Third generation: 162</td>
</tr>
<tr>
<td>Cross River Ferry</td>
<td>Total: 6</td>
<td>53-54</td>
</tr>
<tr>
<td>CityHopper</td>
<td>Total: 3</td>
<td>78</td>
</tr>
</tbody>
</table>

A total of 25 terminals along the Brisbane River are serviced by Brisbane’s ferry system with the CityCat servicing a linear route consisting of 18 terminals. Terminals situated near key destinations include UQ St Lucia for The University of Queensland, QUT Gardens Point for Queensland University of Technology and Southbank for Griffith University and TAFE Queensland Brisbane, in addition to North Quay, Southbank and Riverside which service Brisbane’s financial district and key entertainment and tourist destinations (BCC, 2017). The system is notable for its regular-interval timetable with 7.5 minute headways in peak hour and 15 minute headways in off peak. An express CityCat route also operates during peak times, and this route will skip terminals with lower demand. The remaining 7 terminals are serviced only by the Cross River Ferry and/or the CityHopper.

Travel on CityCats is subsidized by the local government, and all public transport fares in Brisbane are standardised based on travel origin and destination. Smart cards – called Go-Cards in South East Queensland – were introduced to replace paper fare tickets in 2008 and now account for at least 90% of public transportation fare transactions (Soltani et al., 2015).
The Go Card is zone-based, wherein commuters are charged based on the number of zones travelled, as can be seen in Figure 1 above. Commuters are able to use the same card for transit on all public transportation modes. Whilst these smart card systems provide convenience for commuters, they also provide data for transportation research. All Go-Card transactions are geospatially recorded and time-stamped. Furthermore, passengers are required to touch on and touch off when beginning and ending a trip (i.e. when boarding or alighting a bus or ferry, and when entering or leaving a train station). Consequently, origin and destination records are readily available for each trip and thus journey (Soltani et al., 2015).

Figure 1: Map showing linear route of CityCat (blue), CityHopper (red) and Cross River Ferry (yellow) on Brisbane River. (Translink, 2015)

4. Methodology

4.1. Go Card Data Structure, Extraction and Cleaning

For this investigation, analysis was conducted using a one-month slice of Go Card transaction data specific to April 2016. Each valid transaction involving the usage of a Go Card includes the information presented in Table 2 below.

GTFS data publically available from Translink can be used to identify the name and longitude and latitude of the stops. For example, 317590 is Riverside Ferry Terminal whilst 319665 is UQ St Lucia Ferry Terminal.

In its unfiltered form, the data obtained from Translink included transactions for all of South-East Queensland for bus, rail, ferry and light rail. To analyse Go Card data specific to ferry usage, an application was developed with C++ in Microsoft Visual Studio and Qt. Transactions which contained “BCC Ferries” as an operator ID were written to a separate csv file. Furthermore, numerous errors were found wherein a traveller’s boarding stop and alighting stop were identical. These transactions were removed, resulting in a total number of 283,171 ferry trips for April 2016.
### Table 2: A sample transaction of Go Card data

<table>
<thead>
<tr>
<th>Operator ID</th>
<th>BCC Ferries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operations Date</td>
<td>1-Apr-2016</td>
</tr>
<tr>
<td>Route</td>
<td>1</td>
</tr>
<tr>
<td>Service</td>
<td>NHAM-106 UQ St Lucia Ferry Terminal</td>
</tr>
<tr>
<td>Direction</td>
<td>Downstream</td>
</tr>
<tr>
<td>Smartcard ID</td>
<td>10732919</td>
</tr>
<tr>
<td>Scheduled Start</td>
<td>1-Apr-2016 17:24:00</td>
</tr>
<tr>
<td>Vehicle</td>
<td>511</td>
</tr>
<tr>
<td>Boarding Date</td>
<td>1-Apr-2016 17:52:30</td>
</tr>
<tr>
<td>Alighting Date</td>
<td>1-Apr-2016 18:41:25</td>
</tr>
<tr>
<td>Passenger Type</td>
<td>Tertiary Student</td>
</tr>
<tr>
<td>Ticket Type</td>
<td>Go Card</td>
</tr>
<tr>
<td>Boarding Stop</td>
<td>317590</td>
</tr>
<tr>
<td>Alighting Stop</td>
<td>319665</td>
</tr>
<tr>
<td>Journey Number</td>
<td>2016040107455185001189323</td>
</tr>
<tr>
<td>Trip Sequence</td>
<td>2</td>
</tr>
</tbody>
</table>

### 4.2. Developing Visualisations of Ferry Data

Microsoft Excel was used for visualising the output data filtered through C++. In order to investigate terminal demand and performance measured through dwell time, a number of plots were developed which shall be explained in the following sub-sections.

#### 4.2.1. Monthly Variation in Alighting or Boarding Demand with Time of Day

A box and whisker plot was developed to visualise terminal demand for boarding and alighting on weekdays as it varied throughout the day. The minimums, interquartile values and maximums used were deduced from alighting passenger values for a set time interval over all weekdays in April 2016. Weekends were removed from this analysis, as weekend and weekday demands vary significantly. Weekday demands reach distinct maximums during peak travel times whilst weekend demand remains predominantly uniform throughout the day. (Soltani et al., 2015). Additionally, 25-04-2016 was removed as this was a public holiday.

For the example plot Figure 2, the time of day range used is 13:00 – 21:00, with time intervals of 30 minutes. The y-axis shows the number of alighting passengers. With this example, alighting passengers reaches a maximum during afternoon peak at 17:30 – 18:00 as evident from the increasing median value. Meanwhile, the interquartile ranges, minimum and maximum values of alighting demand also peak during this time interval.

#### 4.2.2. Passenger Alighting and Boarding Time Durations upon Vessel Arrival at Ferry Terminal

A plot was developed to facilitate analysis of passenger alighting and boarding behaviour at ferry terminals. Data for a single day was used, this being a Wednesday dated 6-Apr-2016. Every alighting and boarding transaction which occurred on this date for a specific ferry terminal is plotted as a point, however the filtered raw data must first be processed to generate the values required by the axes depicted in Figure 3. The x-axis shows the arrival time of the ferry at a specific terminal. This is determined to be the first occurring transaction for the ferry when it arrives at the terminal, be it for an alighting or boarding passenger. Meanwhile, the y-axis measures the time duration in seconds of passenger transactions which follow the first occurring passenger transaction as shown in Equation 1. Table 3 provides an example of how this data was processed.
Equation 1: Time duration of transaction for dwell time calculation

\[
\text{Time Duration of Transaction } i = \text{Time of Transaction } i - \text{Time of First Occurring Transaction}
\]

Figure 2: Example Box-and-Whisker plot for variation in number of alighting passengers on weekdays with time of day

Figure 3: Sample plot for dwell time duration of each ferry due to passenger alighting and boarding
Table 3: Time duration calculation for ferry arriving at 6:14:31PM.

<table>
<thead>
<tr>
<th>Transaction Number</th>
<th>Time of Transaction</th>
<th>Time of First Occurring Transaction</th>
<th>Time Duration of Transaction (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6:14:31 PM</td>
<td>6:14:31 PM</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>6:14:37 PM</td>
<td>6:14:31 PM</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>6:14:40 PM</td>
<td>6:14:31 PM</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>6:14:45 PM</td>
<td>6:14:31 PM</td>
<td>14</td>
</tr>
<tr>
<td>5</td>
<td>6:14:48 PM</td>
<td>6:14:31 PM</td>
<td>17</td>
</tr>
<tr>
<td>6</td>
<td>6:15:06 PM</td>
<td>6:14:31 PM</td>
<td>35</td>
</tr>
<tr>
<td>7</td>
<td>6:15:13 PM</td>
<td>6:14:31 PM</td>
<td>42</td>
</tr>
<tr>
<td>8</td>
<td>6:15:19 PM</td>
<td>6:14:31 PM</td>
<td>48</td>
</tr>
<tr>
<td>9</td>
<td>6:15:41 PM</td>
<td>6:14:31 PM</td>
<td>70</td>
</tr>
<tr>
<td>10</td>
<td>6:15:59 PM</td>
<td>6:14:31 PM</td>
<td>88</td>
</tr>
<tr>
<td>11</td>
<td>6:16:08 PM</td>
<td>6:14:31 PM</td>
<td>97</td>
</tr>
<tr>
<td>12</td>
<td>6:16:10 PM</td>
<td>6:14:31 PM</td>
<td>99</td>
</tr>
<tr>
<td>13</td>
<td>6:16:13 PM</td>
<td>6:14:31 PM</td>
<td>102</td>
</tr>
<tr>
<td>14</td>
<td>6:16:25 PM</td>
<td>6:14:31 PM</td>
<td>114</td>
</tr>
</tbody>
</table>

Alighting and boarding transactions are represented with different symbols in the plot. When alighting / boarding conventions are adhered to by passengers, then the plot should reflect existing passengers alighting prior to new passengers boarding. However, if there are deviations from this convention, it may be of interest to determine the cause as the CityCat ferry gangway is narrow, and only allows passage in one direction at a time. Additionally, to optimise the ferry dwell time at a terminal, it would be ideal to begin investigating the causes of the delays which exist between transactions for a specific ferry. Abnormal delay can be identified as being large differences in time duration between two transactions. For example, when considering Table 3 and/or Figure 3, relatively large delays occur for transactions 6, 9 and 10, thus decreasing the time efficiency of the ferry at this specific terminal.

The ferry headway can be determined by measuring the time difference between each vertical series of transactions. For this sample, it can be observed that the ferry is reliable with minimal variance between ferry arrival times. However, it may be of benefit to simultaneously acknowledge information from the plots developed in 3.2.1, wherein a large variance in weekday demand is evident. Given a significant variance in passenger demands, fixed schedules may not be the most efficient manner through which travel patterns are serviced (Fu, Liu, & Calamai, 2003). However, if demand data from plots similar to Figure 3 is analysed in real-time, it may be possible to begin acquiring the information necessary for introducing dynamic scheduling, as the real-time demand of a ferry terminal can be used to determine the longest time an arriving passenger should wait at the terminal.

A series of three plots were developed for analysing terminal performance with respect to the terminal's time efficiency in servicing specific quantities of alighting, boarding and total passengers. Similar to the plot developed in 3.2.2, all passenger Go Card transactions are recognised and as such, each data point represents an individual transaction. Additionally, the data was limited to a single day that being Wednesday 6-Apr-2016. The x-axis displays the alighting, boarding or total passenger count whilst the y-axis shows the ferry dwell time required for this number of passenger movements to be completed.

4.2.3. Relationship between Number of Passengers Alighting, Boarding or Total Passengers with Ferry Dwell Time Duration

A series of three plots (Figure 4) were developed for evaluating the terminal's time efficiency in servicing specific quantities of alighting, boarding and total passengers. All passenger smart card transactions are recognized and thus each data point represents an individual
transaction. The x-axis displays the alighting, boarding or total passenger count whilst the y-axis shows the ferry dwell time required for this number of passenger movements to be completed. To identify a potential correlation between passenger flow quantity and dwell time, the mean dwell time was calculated and plotted for each passenger flow quantity. A linear trend line was then fitted to the mean dwell time values. Although linearity for dwell time due to alighting passengers is strong ($R^2 = 0.98$), the coefficient of correlation decreases for boarding passengers ($R^2 = 0.79$) and further for total passengers ($R^2 = 0.51$).

A number of qualitative observations can also be noted. Minimum dwell time due to a set quantity of passenger movements can be identified as a linear pattern. This minimum dwell time value could provide a means of evaluation for each terminal, wherein dwell times due to passenger counts for each ferry arriving can be compared to the minimum possible dwell time. It would also be of interest to determine causes for deviations from this minimum. In some circumstances, e.g. for 15 total passenger movements, the dwell time due to passengers has a minimum of approximately 20s yet may vary to 750% of this value at 150s. Furthermore, there is a decrease in $R^2$ from the alighting plot to the boarding plot. Whereas alighting passengers will generally disembark the vessel upon arrival, boarding passengers may arrive late resulting in the vessel waiting and thus having a longer dwell time. This could be the reason for the greater number of delays observable for the boarding plot.

Figure 4: Sample plot for ferry dwell time duration due to number of passengers alighting or boarding or total movements

![Sample plot for ferry dwell time duration due to number of passengers alighting or boarding or total movements](image)

5. Results and Discussion

The above data visualisations were created for Riverside (Figures 5 & 6), Bulimba (Figures 7 & 8) and Mowbray Park (Figures 9 & 10) Ferry Terminals. Riverside services Brisbane’s major financial district, whilst Bulimba and Mowbray Park are popular terminals which service residential districts. Through these methods, inefficiencies in the systems can be identified. The next step would be to focus on their causes.
5.1. Demand Variation with Time of Day

These plots highlight a significant increase in demand and variance in demand during peak travel times, whilst outside of peak, demand and variance are uniform. Further analysis can then be undertaken to determine the causes for these large variances in terminal usage despite all the days being weekdays. However, for this dataset, these variances could potentially be due to the Easter school holidays which start in March and conclude on the 10th of April in 2016.

Figure 5: Box-and-Whisker plot for variation in number of alighting passengers with time of day for Riverside ferry terminal in April 2016

Figure 6: Box-and-Whisker plot for variation in number of boarding passengers with time of day for Riverside ferry terminal in April 2016

Figure 7: Box-and-Whisker plot for variation in number of alighting passengers with time of day for Bulimba ferry terminal in April 2016
5.2. Alighting and Boarding Time Durations upon Ferry Arrival

The majority of boarding and alighting transactions are successive with little time difference between each transaction. However, Figures 11, 12 and 13 also reveal numerous instances of abnormally large delays for all terminals. These delays predominantly occur due to the movements of boarding passengers, and could potentially be due to passengers who arrive...
late for boarding. Note that a number of these values may be inaccurate due to the commuters touching on / touching off before the ferry docks at the terminal, as the go card reader on-board may be activated before the ferry itself docks.

To allow for more in-depth analysis, the range of arrival times investigated should be decreased. Upon identifying the transactions which are causing the delays, it would also be of interest to investigate the other information provided on the transaction by Go Card data.

Figure 11: Plot for dwell time duration of each ferry due to passenger alighting and boarding at riverside ferry terminal for April 2016

Figure 12: plot for dwell time duration of each ferry due to passenger alighting and boarding at Bulimba ferry terminal for April 2016
Figure 13: Plot for dwell time duration of each ferry due to passenger alighting and boarding at Mowbray park ferry terminal for April 2016

5.3. Dwell Time Duration due to Quantity of Passenger Movements

The series of plots for dwell time duration due to passenger movements also reflects the numerous delays previously depicted in Section 4.2. These delays are particularly evident for Riverside Ferry Terminal, as in Figure 14, where certain ‘strings’ of transactions can be distinguished from the main body of transactions as they required a greater duration of time. It would be of benefit to investigate why some passenger movements for specific ferries are time efficient (i.e. those which are or are close-to the minimum dwell time durations) whilst other passenger movements experience delays and dwell time durations significantly greater than minimum values.

Figure 14: Plot for dwell time duration due to quantity of (a) alighting passengers, (b) boarding passengers and (c) total alighting and boarding passengers for riverside ferry terminal in April 2016
6. Conclusions and Discussion

To successfully deploy sustainable transport strategies, focus must first be placed on understanding commuter behaviour patterns. Geospatial, time-stamped smart card data can offer numerous insights into these behaviours. As such, this paper sought to demonstrate how smart-card data could be used to analyse the performance of ferry terminals for Brisbane’s ferry network. In the process, new visualisations for depicting smart-card data were developed. Subsequently, numerous inefficiencies and delays can easily be identified through observation. Unfortunately, the causes of these issues cannot be established without further analysis into the smart-card data.

It is suggested that future work continue deeper with microscopic analysis, and focus on extracting the transactions which cause inefficiencies and delays. Furthermore, although ferry smart card data was considered for this paper, the same methodologies can be applied to bus and train services. Additionally, this analysis can assist in collecting information on the public transport system structure and operating procedures to specify microscopic simulation model parameters and input probability distributions as well as to understand the performance of the existing system.

7. Acknowledgements

The authors would like to thank Peter Burnton, Kylie Nixon and Kristy Butler at Arup Brisbane for their support and assistance as project champions for this study which was conducted as an undergraduate research project as a part of the joint program between the Icarus Program (www.civil.uq.edu.au/icarus/home), an undergraduate engagement program developed by the School of Civil Engineering at the University of Queensland, and the QUEST Honors Program at the University of Maryland, USA. The authors would also like to thank TranksLink for access to data essential to this study.

8. References


